Trapping Virtual Pores by Crystal Retro-Engineering

Marc A. Little[1], Michael E. Briggs[1], James T. A. Jones[1], Marc Schmidtmann[1], Tom Hasell[1], Samantha Y. Chong[1], Kim E. Jelfs[2], Linjiang Chen[1] and Andrew I. Cooper[1]*

[1]Department of Chemistry and Centre for Materials Discovery, University of Liverpool, L69 3BX, United Kingdom, [2] Department of Chemistry, Imperial College London, SW7 2AZ, United Kingdom, * E-mail: aicooper@liverpool.ac.uk

Stable, guest-free porous molecular crystals are uncommon. By contrast, organic crystals with guest-occupied cavities are frequently observed, but these cavities tend to be unstable and collapse on removal of the guests – this feature has been referred to as "virtual porosity". Here, we have trapped the virtual porosity in an unstable, low-density organic molecular crystal by introducing a second molecule that matches the size and shape of the unstable voids. We call this strategy 'retro-engineering' because it parallels organic retrosynthetic analysis, and it allows the metastable 2-D hexagonal pore structure in an organic solvate to be trapped in a binary cocrystal. Unlike the crystal with "virtual porosity', the cocrystal material remains single crystalline and porous after guest removal by heating.

Macrocycles, molecular cages, and other organic molecules often form crystalline solvates, or inclusion compounds, that comprise guest-filled cavities or channels. These channels are reminiscent of pores, but they are typically not stable to guest removal. Barbour[1] referred to this as "virtual porosity" because it can be created, in a virtual sense, by deleting guests *in silico*. Molecular organic crystals with conventional porosity, where the pores are stable in the absence of guests, are much rarer[2-6]. In principle, solvated crystals with virtual porosity provide a structural blueprint for analogous functional materials with true, conventional porosity. Indeed, "virtual" porosity can sometimes be preserved by the application of careful technique. For example, Mastalerz generated conventional mesoporosity in a large organic cage molecule, but only when the solvent was removed by a specific series of solvent exchanges[7]. However, while the use of techniques such as solvent exchange or supercritical drying[8], can preserve porosity in some cases, the resulting materials may often be too fragile for practical applications. Here, we demonstrate that virtual porosity can be trapped in a stable cocrystal via a crystal retro-engineering approach. This involves identifying a second

molecule with the appropriate shape and dimensions required to stabilize a proportion of the solvent-filled "virtual pores".

Binary, non-covalent cocrystals have been studied extensively[9]. In addition to salts and cocrystals of simple organic molecules, a wide variety of host-guest systems are known, such as the macrocyclic inclusion complexes of cucurbiturils[10] and neutral calixarene analogues[11]. Directionally bonded networks, often involving hydrogen bonding interactions[12], are especially common in binary or higher-order cocrystals. Indeed, complementary hydrogen bond pairing has been used to predict probable cocrystals[13,14], or, as described by Desiraju, to 'synthesise' molecular crystals[15]. However, the structures of cocrystals are still difficult to manipulate in a programmed way. Isomorphous substitutions in molecular crystals are uncommon because even a small change in the structure of the molecular building blocks can lead to a large change in the resulting crystal packing. Hence, there are no real molecular analogues of isoreticular metal-organic frameworks (MOFs)[16,17], where families of isostructural porous materials exist for a wide range of organic linkers.

Despite these difficulties, a number of crystal engineering strategies have been developed for organic cocrystals. For example, Desiraju showed that a weakly bound bipyridine molecule could be replaced in a ternary molecular solid using five suitably-sized analogues without altering the crystal packing[18]. Likewise, Davis found that the structure of a molecular scaffold composed from a steroidal urea contains a 1-D channel that can be decorated with a number of chemical functionalities without altering the overall packing[19-21]. Of particular relevance here, McKeown stabilized an unstable solvated pthalocyanine crystal by using a ligand exchange approach to produce molecular cocrystals that could then be desolvated to produce conventional porosity[22]. This 'wall tying' report is elegant but so far unique, and it relies on the specific chemistry of metal pthalocyanines.

We have produced porous molecular crystals in a modular way by cocrystallizing two[23,24] or even three[25] organic cage molecules. However, this has been limited to structurally similar molecules, and our 3-component cocrystals[25] comprised three [4+6] imine cages, **CC1**, **CC3** and **CC4**, each with the same tetrahedral symmetry and approximate size (Figure 1).

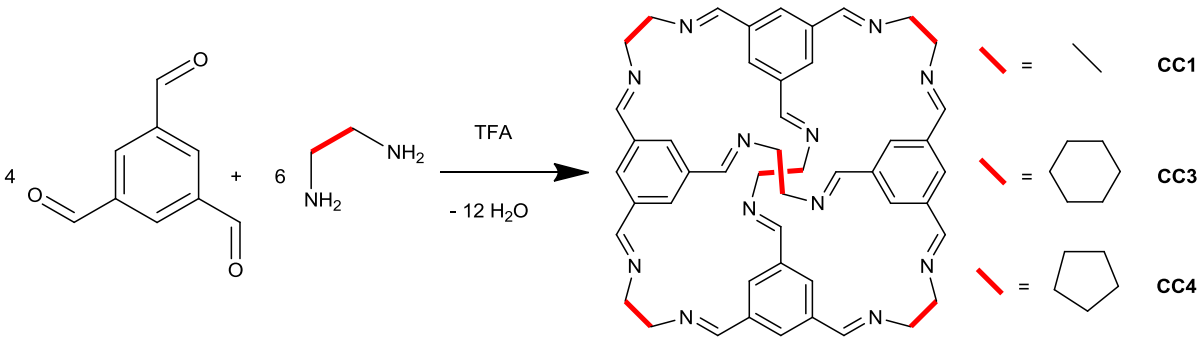


Figure 1: Synthesis of 'porous organic cage' molecules: Reaction scheme for preparation of tetrahedral [4+6] cage molecules **CC1**, **CC3** and **CC4**, catalysed by trifluoroacetic acid (TFA).

As these various examples show, programmable molecular substitution in organic cocrystals is still highly challenging, and most examples involve relatively small structural changes in the constituent molecular building units, as in our ternary 'porous organic alloys'[25].

Here, we show that porous molecular cocrystals can be formed from two cages with dissimilar size and symmetry: a tetrahedral [4+6] cage and a smaller trigonal [3+2] cage. The resulting porous cocrystal was 'retro-engineered' from the closely related structure of a metastable solvate of the [4+6] cage, which has 2-D "virtual porosity" that is not stable to solvent guest removal. Specifically, the [3+2] partner cage fills an unstable cavity in the virtually porous solvate that would otherwise collapse on removal of the solvent. Incorporation of the [3+2] partner cage results in a stable, permanently porous cocrystal with the same underlying 2-D pore structure as that of the [4+6] molecular crystal.

Results

CC3-*R* is a homochiral, shape-persistent porous organic cage prepared by the reaction of triformylbenzene with (*R*,*R*)-1,2-cyclohexanediamine (Figure 1)[26]. When crystallized from neat dichloromethane (CH_2Cl_2) or chloroform ($CHCl_3$), **CC3**-*R* packs in a window-to-window arrangement, thus generating an interconnected, diamondoid 3-D pore network running through the intrinsic cage voids. This polymorph, **CC3**α, (Figure 2i), was shown by crystal structure prediction (CSP) calculations[23] and by DFT calculations for cage dimers[24] to be the most stable crystal packing for **CC3**-*R*.

While **CC3**α is the most stable desolvated polymorph for **CC3**-*R*, we have found recently that this cage can also form metastable solvates with different crystal packings, for example when **CC3** is crystallized using diethyl ether[27]. Here, crystallization of **CC3**-*R* from a $CHCl_3$/MeOH solution (< 50 % v/v $CHCl_3$) affords two structurally similar solvate phases as a concomitant mixture. Both solvates transform to the same desolvated phase upon heating to 300 K, and hence only one of these solvate forms is described here, **phase 1** (Supplementary Information, section 2.0–2.4, Figure S1–11). The MeOH solvate of **CC3**-*R*, referred to as **phase 1**, was refined in the chiral monoclinic space group *I*2, with one complete **CC3**-*R* molecule and a number of partially-occupied solvent molecules in the asymmetric unit. In the crystal packing of **CC3**-*R*·$(MeOH)_{12}$·$(H_2O)_4$ (**phase 1**), each cage molecule packs in a window-to-window arrangement with three neighboring cage molecules at a cage centroid separation distance of ~ 11.7 Å (Figure 2a). Extension of this window-window pairing arrangement in a 2-D fashion generates an interconnected, solvent-filled "virtual" pore network running through the intrinsic cage voids (Figure 2b). These 2-D honeycomb pore networks are approximately planar and are layered in an offset manner with interlayer voids between the 2-D networks (Figure 2c). The volume of the unit cell per cage molecule is 500 $Å^3$ larger for **phase 1** compared to the thermodynamic, solvent-free polymorph, **CC3**α (Supplementary Information, section 2.5, Table S1), indicating less efficient packing of the cages in the presence of MeOH. Most of the additional volume is located in intralayer extrinsic voids between the hexagonally arranged **CC3**-*R* molecules.

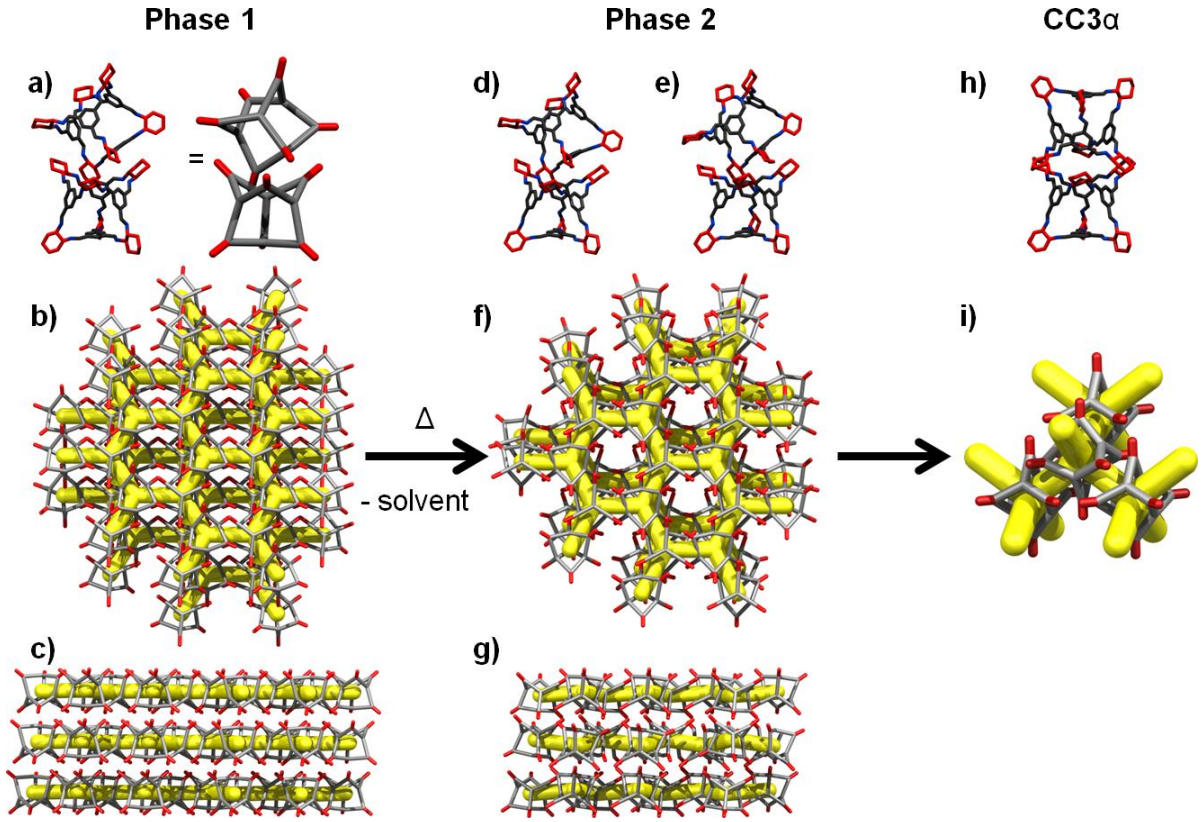


Figure 2: A metastable solvate phase of a porous cage, phase 1, has "virtual" 2-D pores but this transforms over time to a denser phase, CC3α, with 3-D pores: a, Representation of the single crystal structure for **CC3**-*R*·$(MeOH)_{12.5}$·$(H_2O)_3$ (**phase 1**) showing window-to-window cage pairing arrangement. **b**, Hexagonally-arranged **CC3**-*R* molecules generate a 2-D honeycomb, solvent-filled "virtual pore" network (yellow), shown in perspective view [101], and **c**, [010]. **d**, **e**, Single crystal structure after heating to remove solvent to generate (**CC3**-*R*)$_2$ (**phase 2**) showing cage window pairings from crystallographically independent pore networks. **f**, 2-D honeycomb pore network in perspective view [001], and **g**, [010]. **Phase 2** is metastable and transforms easily to the thermodynamic, solvent-free polymorph, **CC3**α, with a 3-D pore network, **h** and **i**.

To investigate the stability of **phase 1** to solvent removal, a flash-frozen single crystal was gradually heated from 100 K to 300 K in 10 K intervals at a ramp rate of 5 K min^{-1}; the equilibration time at each temperature was 30 minutes. This careful, stepwise process prevents loss of singularity upon desolvation and results in a single-crystal-to-single-crystal transformation to an alternate monoclinic *C*-centred unit cell. This new phase, **phase 2**, was refined in the space group *C*2 with two complete **CC3**-*R* molecules in the asymmetric unit. No solvent molecules were assigned in the single crystal structure of **phase 2**, although it is possible that a small amount of diffuse solvent still occupies the interconnected voids (Supplementary Information, section 2.2). Comparing **phase 1** with **phase 2** reveals a

rearrangement in molecular packing, but no significant change to the structure of the cage itself or the intralayer voids between the hexagonally arranged **CC3**-*R* molecules (*c.f.*, Figure 2b & 2f). As for **phase 1**, cage window pairing in **phase 2** leads to alignment with three neighbouring cage molecules (Figure 2d & 2e), and extending this in 2-D generates an offset honeycomb pore network (Figure 2f). However, unlike the "virtual pores" in **phase 1**, the 2-D pore layers in **phase 2** are puckered, rather than planar (*c.f.*, Figure 2c & 2g) and the offset layering for the 2-D networks of **CC3**-*R* molecules in **phase 2** is laterally sheared from that found in **phase 1** (*c.f.*, Figure 2b & 2f). For **phase 2** the 2-D networks also pack ~ 0.7 Å closer together than for **phase 1** (Supplementary Information, section 2.5, Table S1); a structural transformation which reduces the interlayer lattice voids. The transformation to **phase 2** reduces the volume of the unit cell per cage molecule by 179 $Å^3$ with respect to **phase 1** (Supplementary Information, Table S1).

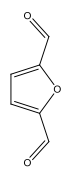
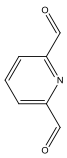
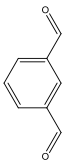
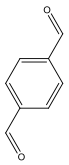
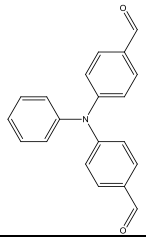
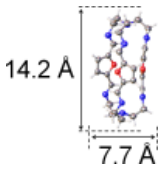
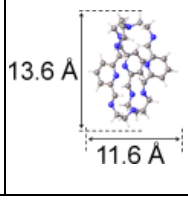
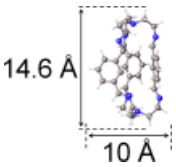
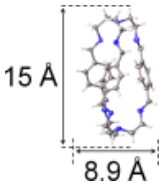
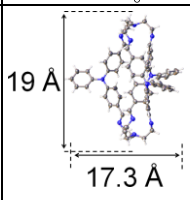
Phase 2 persists up to 400 K (Supplementary Information, section 2.5, Table S1), although heating to this temperature causes a complete loss of crystal singularity after 3 hours. We also investigated this transformation for bulk phases, screening a number of samples with varying concentrations of MeOH in $CHCl_3$, or MeOH in CH_2Cl_2 (Supplementary Information, section 2.6, Table S2). Two solvated phases of **CC3**-*R* were isolated from CH_2Cl_2/MeOH solvent mixtures, which were isostructural to those isolated from $CHCl_3$/MeOH (Supplementary Information, section 2.7–2.8, Figure S12–13). Analysis of powder XRD (PXRD) patterns recorded for these bulk samples after full evacuation of the pores revealed only one detectable phase: the thermodynamic polymorph, **CC3**α (Figure 2i; Supplementary Information, Section 2.9–2.11, Figure S14-21). Hence, unless very careful steps are taken during desolvation, thermal transformation of **phase 1** to bulk **CC3**α occurs via metastable **phase 2** (Figure 2b, 2f & 2i). This transformation increases the crystallographic density from 0.836 g/cm^3 in **phase 2** to 0.973 g/cm^3 in **CC3**α. Previously, Atwood reported a porous frustrated crystal packing of the host molecule *p-tert*-butylcalix[5]arene that could be trapped by carefully activating a toluene solvate crystal at 120 °C while activation at 160 °C resulted in a transformation of the host molecule to its thermodynamic, non-porous polymorph[28]. In our case, the metastable structure, **phase 2**, could not be isolated on bulk scale.

Phase 1 and **phase 2** are unstable with respect to **CC3**α because of the additional extrinsic voids between the cages. The extrinsic intralayer hexagonal voids in **phase 1** (Figure 3a) and **phase 2** (Figure 3b) are capped by **CC3**-*R* molecules layered directly above and below. The resulting enclosed voids can be described by vertical and horizontal vectors (Supplementary

Information, Figure S22–23) and by their shape (Figure 3c). In principle, the 2-D virtual intralayer pores in **phase 1** and **phase 2** might be preserved by identifying an involatile molecule that fits precisely in these intralayer extrinsic voids (cyan spheres, Figure 3a & b), hence preventing structural reorganization to **CC3**α. That is, a stable cocrystal with conventional 2-D porosity might be 'retro-engineered' from the unstable, virtually porous **phase 1** solvate.

We noticed that the symmetry and dimensions of these intralayer extrinsic voids are reminiscent of certain trigonal [3+2] cycloimine cages (cages **[3+2]**$_{\#1\text{-}5}$, Table 1, *c.f.*, Figure 3c & 3d), synthesized by cycloimine condensation reactions of tris(2-aminoethyl)amine with different dialdehydes (Supplementary Information, section 3.0, Table S3, Figure S24–34). Cages **[3+2]**$_{\#1\text{-}4}$ were synthesized via a modified literature method[29,30]. Cage **[3+2]**$_{\#5}$ is a novel macrocycle, reported here for the first time.

Table 1. Structures of trigonal [3+2] organic cages prepared by reaction of dialdehyde precursors with tris(2-aminoethyl)amine.

	[3+2]$_{\#1}$	**[3+2]**$_{\#2}$	**[3+2]**$_{\#3}$	Reduced **[3+2]**$_{\#4}$	**[3+2]**$_{\#5}$
Chemical structure of dialdehyde					
Structure of trigonal [3+2] cage molecules					

Molecular structures from single crystal structures are given for three of the five cages (**[3+2]**$_{\#1,4,5}$). Ellipsoids displayed at 50 % probability level, solvent is omitted. X-ray structures were not obtained for cages **[3+2]**$_{\#2,3}$, so molecular models were constructed and a short conformer search (1000 steps in a low-mode search in MacroModel with the OPLS-AA forcefield) was performed to find the lowest energy conformation[31]. Approximate cage dimensions calculated using distant hydrogen–hydrogen contacts plus Van der Waals radius shown for comparison (molecules not shown on common scale).

We screened the propensity of these [3+2] cages to cocrystallize with **CC3**-*R*, and thus to direct **CC3**-*R* to be isostructural with **phase 1** by filling the unstable intralayer extrinsic voids. A cocrystallization screen was carried out where the **[3+2]**$_{\#1\text{-}5}$ cages were mixed, individually, with **CC3**-*R* in a 1:2 molar ratio. This ratio was chosen because we expected that only a single trigonal [3+2] cage could be accommodated in each intralayer extrinsic void, and there are half as many intralayer extrinsic voids in the crystal lattice in relation to the number of **CC3**-*R* molecules in **phase 1** and **phase 2**. In each case, the two different cages were dissolved in $CHCl_3$, followed by crystallization, either by slow solvent evaporation or by vapor diffusion of an antisolvent (MeOH). Using these conditions, only one of the trigonal cage molecules, **[3+2]**$_{\#5}$, led to a new phase that contained **CC3**-*R*.

The crystallization of **CC3**-*R* and **[3+2]**$_{\#5}$ by antisolvent diffusion resulted in triangular, plate-like crystals (Supplementary Information, section 4.0, Figure S35), as opposed to the needle shaped crystal habit of **phase 1** or the octahedral crystal habit known for **CC3**α[32]. Structural identification of this phase by single crystal XRD revealed the formation of a binary cocrystal of the desired composition, (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$, which had crystallised in the chiral trigonal space group *R*32 with the trigonal **[3+2]**$_{\#5}$ molecules positioned in the intralayer extrinsic voids. We are not aware of a previous example of a binary cocrystal comprising two structurally dissimilar organic cages[23-25]. The asymmetric unit comprises one third of a **CC3**-*R* molecule centered around a threefold rotation axis, plus one sixth of a **[3+2]**$_{\#5}$ cage centered around the intersection point of three twofold rotation axes and one threefold rotation axis. The **[3+2]**$_{\#5}$ cage is disordered over two positions, in which it is rotated 43° on a threefold rotation axis that runs the length of the molecule, passing through the tertiary amine 'poles'. One $CHCl_3$ molecule was found in the intrinsic **CC3**-*R* cavity along with, in total, two disordered MeOH solvent molecules located in an extrinsic void spanning between two window-to-window packed **CC3**-*R* molecules.

The solvated single crystal, (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$, was heated *in situ* under a dry nitrogen gas stream (Supplementary Information, section 4, Figure S36–43). Heating to 300 K caused partial loss of the $CHCl_3$ and MeOH solvent; subsequent heating to 450 K led to complete loss of lattice solvent without any evidence of crystal degradation or change in structure. An improved data set was obtained on cooling the single crystal to 100 K, and the difference map again indicated the absence of any electron density that could be ascribed to solvent.

The crystal packing of (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ closely resembles that of **phase 1** and **phase 2** in terms of the orientation of the **CC3**-*R* molecules, with the addition of the smaller trigonal cage, **[3+2]**$_{\#5}$, in the extrinsic voids (Figure 3). Indeed, the distance between the cage centroids of the hexagonally arrayed **CC3**-*R* molecules is 12 Å, only slightly longer than the equivalent distance in **phase 1** (11.7 Å) or **phase 2** (range 11.3–11.8 Å). In addition, crystal structure comparison between **phase 2** and (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$, reveals that these crystal structures are isostructural with respect to the 2-D packing of **CC3**-*R* molecules (Figure 4). Hence, an equivalent 2-D honeycomb pore network exists in desolvated (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ (Figure 3h & 3i). The three terminal phenyl rings of the **[3+2]**$_{\#5}$ cage impinge on the pore channel that runs through intrinsic **CC3**-*R* cavities (Figure 3h), therefore modifying the limiting diameter of these pores. The vertical dimension of the **[3+2]**$_{\#5}$ cage molecules is also longer than the equivalent dimension for **CC3**-*R*. As a result, the trigonal [3+2] cages protrude out of the 2-D pore network (Figure 3i), and these [3+2] cages act as 'pegs' that confer additional stability, preventing lateral shearing of the 2-D layers upon desolvation (Figure 3k). The trigonal [3+2] cage is size-excluded from the intrinsic **CC3**-*R* cavities: hence, **[3+2]**$_{\#5}$ is able to stabilize the extrinsic voids in **phase 1** / **phase 2** without also filling the 2-D layered pore channels.

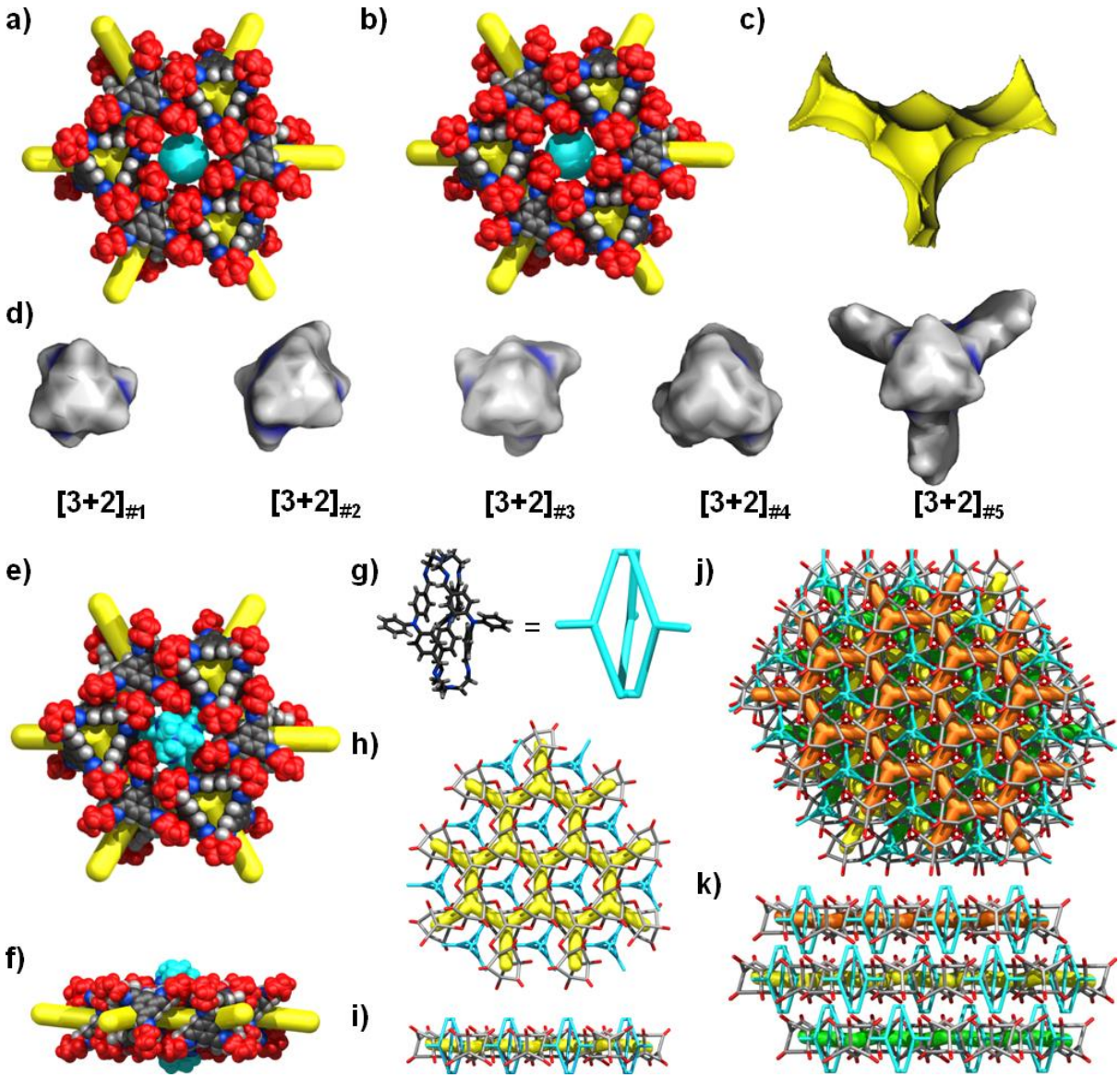


Figure 3: '**Retro-engineering' a binary porous cocrystal: a**, Cyan sphere (radius 4 Å) highlighting the extrinsic void in **phase 1** and, **b**, **phase 2** (Grey, C; blue, N; white, H; except for the cyclohexane vertices, shown in red: 2D pore network shown in yellow). **c**, Visualization of solvent-accessible surface for the extrinsic void in **phase 2** (cyan spheres, **a** + **b**) for a N_2 probe radius (1.82 Å)[33]. **d**, Molecular surfaces for five different trigonal **[3+2]**$_{\#1\text{-}5}$ cage molecules (see Table 1) (PyMol Molecular Graphics System); all molecules shown on same scale. **e**, **f**, Single crystal structure for solvent-free (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$. The carbon atoms of the **[3+2]**$_{\#5}$ cage, which stabilizes the voids in **phase 2**, are colored cyan. **e**, Structure shown in perspective view [001], and, **f**, [010]. The **[3+2]**$_{\#5}$ cage, **g**, intercalates into the crystal structure and stabilizes the 2-D hexagonal pore network, **h**, **i**. Extended crystal lattice showing three layers of offset 2-D pore networks. **j**, **k**; the layered 2-D pore channels are highlighted in orange, yellow, and green. The structures are shown in perspective view [001] (**e**,**h**,**j**) and [010] (**f**,**i**,**k**).

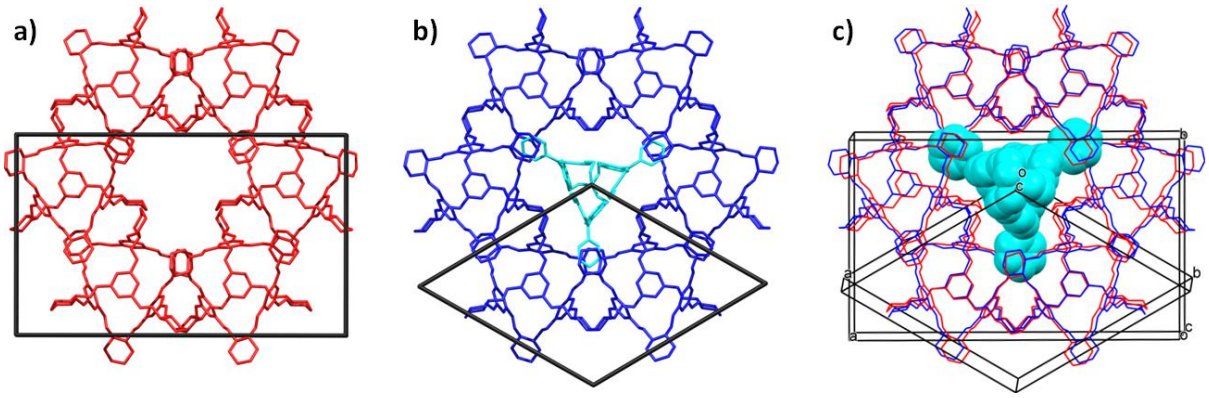


Figure 4: The 2-D porosity in the stable binary cocrystal is 'retro-engineered' from the unstable CC3 solvate: a, Crystal structure of **CC3 phase 2** (red) and, **b**, (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ (blue). **c**, Overlay of crystal structures shows that the 2-D pore layers are isostructural in these two materials. **[3+2]**$_{\#5}$ component shown in cyan; unit cell axes are shown.

One limitation of using a $CHCl_3$/MeOH antisolvent mixture for the preparation of this cocrystal is the simultaneous formation of **phase 1** in which the smaller trigonal [3+2] cage is not included. To address this, a number of alternative cage concentrations and relative stoichiometries were investigated, as was the use of alternative antisolvents (Supplementary Information, Table S4). After some optimization, we found that dissolving the two cages in CH_2Cl_2 (instead of $CHCl_3$) and then adding a five-fold volumetric excess of Et_2O led to phase pure crystals of the binary (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ cocrystal after slow evaporation of two thirds of the solvent volume (Supplementary Information, section 5.0). These crystals were activated by heating at 573 K under dynamic vacuum for 16 hours, after which time the crystals were still single, enabling accurate structure solution of a crystal mounted in an environmental gas cell kept under dynamic vacuum (Supplementary Information, Section 5.1, Figure S45). PXRD data, recorded on the same batch of bulk material demonstrated that there is only one crystalline phase present after activation, (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ (Supplementary Information, section 5.2, see Figure S46 for Le Bail fit). The long-term thermal stability of (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ at 573 K is very different from isostructural **phase 1**, which quickly transforms to **CC3**α, via **phase 2**, with only modest heating. This validates the 'retro-engineering' strategy of filling an unstable solvent void with a second, non-volatile cage molecule.

Mapping of the solvent-accessible surface in (**CC3**-*R*)$_2$·**[3+2]**$_{\#5}$ reveals that the intrinsic cage voids are formally disconnected to a probe radius of 1.20 Å due to the insertion of the pendant aromatic rings of **[3+2]**$_{\#5}$ into the pore channels. However, as a result of cooperative diffusion, as observed for formally non-porous calixarenes[34] the cage voids are accessible to

small guests. For example, crystals of (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$** were found to adsorb CO_2 (1.7 mmol/g) with good ideal selectivity over N_2 at 273 K (Supplementary Information, section 6, Figure S48-52). We and others have reported CO_2/N_2 or CO_2/CH_4 selectivity for porous organic cages[35-39]. For (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$** at 1 bar pressure and at 293 K, the ideal CO_2/N_2 selectivity, *S*, is 10. This is lower than the quoted ideal selectivities reported by Zhang *et al*[35-36], but our material has a significantly higher absolute CO_2 uptake (1.0 mmol g^{-1} for (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$** *versus* 0.1–0.25 mmol g^{-1} at 1 bar and 293 K for the Zhang material), although our material has a lower absolute CO_2 uptake than that reported by Mastalerz *et al* (3.3 mmol g^{-1} at 298 K; CO_2/N_2 selectivities not reported).[37].

Discussion

We have successfully introduced functional, gas-selective porosity into a binary cocrystal that was retro-engineered from an unstable solvate that has only "virtual porosity". The cocrystal is exceptionally stable, remaining single crystalline up to 573 K. The approach relies on the close shape match between **[3+2]$_{\#5}$** cage and the unstable voids in **phase 1**, though preliminary experiments suggest a degree of structural tolerance here: for example, the slightly bulkier tolyl analogue of **[3+2]$_{\#5}$** appears to act in much the same way.

In principle, this retro-engineering strategy might be applied to other unstable solvates, providing that an involatile molecule, or crystal coformer, can be identified with the correct size and shape to stabilize the solvent-filled voids (*c.f.*, Figure 3c & 3d, right). This strategy might also allow other solid-state properties to be retro-engineered, transforming the properties of the resulting material, just as structural mimicry has been used previously for photoreactive materials[40]. Unlike McKeown's "wall tying" strategy[22], our retro-engineering approach does not rely on specific metal-ligand chemistry: in principle, therefore, this method might be applicable to almost any molecular crystal where an unstable solvent void can be back-filled by an appropriate involatile molecule, or 'coformer'.

To generalize the approach, however, will require a more directed strategy for identifying appropriate crystal coformers. In this first example, the structural 'retro-engineering' is heuristic, and is based on somewhat qualitative comparisons of the size and geometry of the extrinsic voids with known trigonal [3+2] cages (Figure 3d). In the future, this heuristic approach might be translated into a more quantitative, computational strategy. For example, in the discovery of new drug molecules, protein docking simulations can be used to screen the affinity of a potential target compound for a particular binding site[41]. Likewise, new

zeolite templates have been discovered by computationally 'growing' hypothetical template molecules inside the desired zeolite cavity, and then ranking their stabilizing influence[42,43]. Equivalent strategies might be devised for predicting molecules, or coformers, that confer stability on unstable, low-density organic solvates, such as **phase 1**, which are commonly observed for cages, macrocycles, and other molecules. This could be done, for example, by docking candidate coformers obtained from crystallographic database searches with specific virtual pores in a solvate, thus providing a more general crystal retro-engineering methodology. This could be useful for finding new porous solids, and also for creating other solid-state functions in molecular crystals. For example, crystal retro-engineering might be used to 'trap out' a specific molecular packing observed in an unstable solvate phase that is thought to be useful in organic electronics. While our example here is based on two organic molecules, there is no reason that the method cannot also be translated to organometallic molecules. Crystal retro-engineering might therefore form a useful supramolecular adjunct to synthetic approaches that seek to engineer functionality into molecular crystals via covalent modifications of the constituent molecules.

More ambitiously, crystal structure prediction (CSP) methods[23,44] always reveal a large number of hypothetical crystal packings that are unstable with respect to the global minimum structure, and these are often also lower in density. As CSP methods advance in scope, complexity, and speed, it might become possible to select such 'virtual phases' on the basis of a particular pore topology, or other property of interest (*e.g.*, intermolecular distance, symmetry, polarization, etc.). Crystal retro-engineering could then be used to realise these hypothetical phases in the laboratory via computational selection of an appropriate coformer.

Methods

Materials. 1,3,5-Triformylbenzene was purchased from Manchester Organics, UK. All other chemicals were purchased from Sigma-Aldrich and used as received, unless otherwise stated. **CC3**-*R* was prepared according to methods described previously[26]. **[3+2]$_{\#1\text{-}4}$** were synthesized via modified methods described previously[29,30].

X-ray crystallography. Single crystal XRD data were measured on a Rigaku MicroMax-007 HF rotating anode diffractometer (Mo-Kα radiation, λ = 0.71073 Å, Kappa 4-circle goniometer, Rigaku Saturn724+ detector). Or for (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$**, when collected under dynamic vacuum recorded in an environmental gas cell at beamline I19, Diamond Light Source, Didcot, UK using silicon double crystal monochromated radiation (λ = 0.6889 Å)[45].

Empirical absorption corrections using equivalent reflections were performed with the program SADABS[46]. Structures were solved with SHELXD[47], or by direct methods using SHELXS[47], and reined by full-matrix least squares on F^2 by SHELXL[47]. For (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$** high-resolution PXRD data was collected on a sample loaded in a glass capillary kept under dynamic vacuum using the Mythen-II position sensitive detector (PSD) at beamline I11, Diamond Light Source, Didcot in transmission geometry (λ = 0.82699 Å)[48,49]. The θ circle was rocked through ±15° to improve powder averaging. Analysis of this powder diffraction pattern was carried out using TOPAS-Academic software[50]. Supplementary single crystal XRD data, including structure factors, is available free of charge from the Cambridge Crystallographic Data Centre (CCDC) via www.ccdc.cam.ac.uk/data_request/cif. **CC3**-*R*·(MeOH)$_{11}$·(H$_2$O)$_4$ (**phase 1**) CCDC # 979940. Formula $C_{83}H_{137}N_{12}O_{15}$; M = 1543.05 g·mol^{-1}; monoclinic space group $I2$, colourless crystal; a = 21.6440(9) Å, b = 19.7822(8) Å, c = 23.270(2) Å; β = 105.704(1)°; V = 9591.3(8) Å^3; Z = 4; ρ = 1.069 g·cm^{-3}; μ = 0.074 mm^{-3}; F (000) = 3356; crystal size = 0.42 x 0.42 x 0.39 mm^3; T = 100(2) K; 80290 reflections measured (1.14 < Θ < 23.27°), 13676 unique (R_{int} = 0.0721), 12573 ($I > 2\sigma(I)$); R_1 = 0.1152 for observed; wR_2 = 0.3387 for all reflections; max/min residual electron density = 0.849 and -0.652 e·Å^{-3}; data/restraints/parameters = 13676/12/923; GOF = 1.600. (**CC3**-*R*)$_2$ (**phase 2**) CCDC # 979941. Formula $C_{72}H_{84}N_{12}$; M = 1117.51 g·mol^{-1}; monoclinic space group $C2$, colourless crystal; a = 33.883(8) Å; b = 20.467(5) Å, c = 25.712(6) Å; β = 95.249(8)°; V = 17756(8) Å^3; Z = 8; ρ = 0.836 g·cm^{-3}; μ = 0.050 mm^{-3}; F (000) = 4800; crystal size = 0.42 × 0.42 × 0.39 mm^3; T = 300(2) K; 49205 reflections measured (1.16 < Θ < 19.59°), 15202 unique (R_{int} = 0.0659), 11751 ($I > 2\sigma(I)$); R_1 = 0.0907 for observed; wR_2 = 0.2673 for all reflections; max/min residual electron density = 0.749 and -0.462 e·Å^{-3}; data/restraints/parameters = 15202/0/1513; GOF = 1.075. (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$**·(CHCl$_3$)$_2$·(MeOH)$_8$ CCDC # 979937. Formula $C_{226}H_{268}N_{35}Cl_6O_8$; M = 3815.45 g·mol^{-1}; trigonal space group $R32$, colourless crystal; a = 20.540(1) Å; c = 44.747(3) Å; V = 16349(2) Å^3; Z = 3; ρ = 1.163 g·cm^{-3}; μ = 0.143 mm^{-3}; F (000) = 6105; crystal size = 0.24 × 0.22 × 0.10 mm^3; T = 100(2) K; 93085 reflections measured (1.98 < Θ < 26.37°), 7399 unique (R_{int} = 0.0572), 6528 ($I > 2\sigma(I)$); R_1 = 0.0662 for observed; wR_2 = 0.2001 for all reflections; max/min residual electron density = 0.456 and -0.550 e·Å^{-3}; data/restraints/parameters = 7399/2/541; GOF = 1.062. (**CC3**-*R*)$_2$·**[3+2]$_{\#5}$** recorded under dynamic vacuum CCDC # 979936. Formula $C_{216}H_{237}N_{35}$; M = 3323.41 g·mol^{-1}; trigonal space group $R32$, colourless crystal; a = 20.607(1) Å; c = 44.949(3) Å; V = 16530(2) Å^3; Z =

3; ρ = 1.002 g·cm^{-3}; μ = 0.057 mm^{-3}; F (000) = 5334; crystal size = 0.21 x 0.21 x 0.18 mm^{3}; T = 293(2) K; 47559 reflections measured (1.41 < Θ < 25.50°), 5334 unique (R_{int} = 0.1180), 5185 ($I > 2\sigma(I)$); R_1 = 0.0583 for observed; wR_2 = 0.1801 for all reflections; max/min residual electron density = 0.219 and -0.221 e·Å^{-3}; data/restraints/parameters = 7538/0/463; GOF = 1.036.

Acknowledgements

The authors gratefully acknowledge the Engineering and Physical Sciences Research Council (EP/H000925/1) and European Research Council under the European Union's Seventh Framework Programme/ERC Grant Agreement no. [321156] for financial support. K. E. J. is a Royal Society URF. We thank Rob Clowes for assistance with sorption measurements and Sean Higgins for assistance with robotic handling apparatus. The authors thank Diamond Light Source for access to beamlines I19 (MT8728) and I11 (EE9282) that contributed to the results presented here also Mark R. Warren and Sarah A. Barnett for their assistance during the single crystal gas cell studies. We thank Referee 2 for suggesting an improved version of Figure 4.

Author Contributions

A.C. conceived the project. M.B., M.L. & T.H. prepared the cage molecules. M.L., T.H. & J.J. crystallised and cocrystallised the cage molecules. M.L., M.S. & S.C. interpreted crystal data. M.L., T.H., & L.C. interpreted sorption data. K.J. modelled cage conformers. All authors interpreted structures and contributed to the preparation of the manuscript.

References

1 Barbour, L. J. Crystal porosity and the burden of proof. *Chem. Commun.* 1163-1168 (2006).

2 Zhang, G. & Mastalerz, M. Organic cage compounds - from shape-persistency to function. *Chem. Soc. Rev.* **43**, 1934-1947 (2014).

3 Tian, J., Thallapally, P. K. & McGrail, B. P. Porous organic molecular materials. *CrystEngComm* **14**, 1909-1919 (2012).

4 Holst, J. R., Trewin, A. & Cooper, A. I. Porous organic molecules. *Nat. Chem.* **2**, 915-920 (2010).

5 Mastalerz, M. Shape-persistent organic cage compounds by dynamic covalent bond formation. *Angew. Chem. Int. Ed.* **49**, 5042-5053 (2010).

6 Mastalerz, M. & Oppel, I. M. Rational construction of an extrinsic porous molecular crystal with an extraordinary high specific surface area. *Angew. Chem. Int. Ed.* **51**, 5252-5255 (2012).

7 Zhang, G., Presly, O., White, F., Oppel, I. M. & Mastalerz, M. A permanent mesoporous organic cage with an exceptionally high surface area. *Angew. Chem. Int. Ed.* **53**, 1516-1520 (2014).

8 Nelson, A. P., Farha, O. K., Mulfort, K. L. & Hupp, J. T. Supercritical processing as a route to high internal surface areas and permanent microporosity in metal−organic framework materials. *J. Am. Chem. Soc.* **131**, 458-460 (2008).

9 Herbstein, F. H. Crystalline molecular complexes and compounds: structures and principles. *Oxford University Press, Oxford, United Kingdom* (2006).

10 Day, A. I. *et al.* A cucurbituril-based gyroscane: a new supramolecular form. *Angew. Chem. Int. Ed.* **41**, 275-277 (2002).

11 Danylyuk, O. & Suwinska, K. Solid-state interactions of calixarenes with biorelevant molecules. *Chem. Commun.* 5799-5813 (2009).

12 Aakeroy, C. B. & Salmon, D. J. Building co-crystals with molecular sense and supramolecular sensibility. *CrystEngComm* **7**, 439-448 (2005).

13 Musumeci, D., Hunter, C. A., Prohens, R., Scuderi, S. & McCabe, J. F. Virtual cocrystal screening. *Chem. Sci.* **2**, 883-890 (2011).

14 Grecu, T. *et al.* Virtual screening identifies new cocrystals of nalidixic acid. *Crys. Growth Des.* **14**, 1749-1755 (2014).

15 Desiraju, G. R. Crystal engineering: A holistic view. *Angew. Chem. Int. Ed.* **46**, 8342 – 8356 (2007)

16 Deng, H. *et al.* Multiple functional groups of varying ratios in metal-organic frameworks. *Science* **327**, 846-850 (2010).

17 Eddaoudi, M. *et al.* Systematic design of pore size and functionality in isoreticular mofs and their application in methane storage. *Science* **295**, 469-472 (2002).

18 Tothadi, S., Mukherjee, A. & Desiraju, G. R. Shape and size mimicry in the design of ternary molecular solids: towards a robust strategy for crystal engineering. *Chem. Commun.* **47**, 12080-12082 (2011).

19 Natarajan, R. *et al.* Nanoporous organic alloys. *Angew. Chem. Int. Ed.* **50**, 11386-11390 (2011).

20 Natarajan, R. *et al.* Tunable porous organic crystals: structural scope and adsorption properties of nanoporous steroidal ureas. *J. Am. Chem. Soc.* **135**, 16912-16925 (2013).

21 Travaglini, L., Bridgland, L. N. & Davis, A. P. Cholanamide components for organic alloys; expanding the scope of nanoporous steroidal ureas. *Chem. Commun.* **50**, 4803-4805 (2014).

22 Bezzu, C. G., Helliwell, M., Warren, J. E., Allan, D. R. & McKeown, N. B. Heme-like coordination chemistry within nanoporous molecular crystals. *Science* **327**, 1627-1630 (2010).

23 Jones, J. T. A. *et al.* Modular and predictable assembly of porous organic molecular crystals. *Nature* **474**, 367-371 (2011).

24 Hasell, T., Chong, S. Y., Jelfs, K. E., Adams, D. J. & Cooper, A. I. Porous organic cage nanocrystals by solution mixing. *J. Am. Chem. Soc.* **134**, 588-598 (2011).

25 Hasell, T., Chong, S. Y., Schmidtmann, M., Adams, D. J. & Cooper, A. I. Porous organic alloys. *Angew. Chem. Int. Ed.* **51**, 7154-7157 (2012).

26 Tozawa, T. *et al.* Porous organic cages. *Nat. Mater.* **8**, 973-978 (2009).

27 Little, M. A., Chong, S. Y., Schmidtmann, M., Hasell, T., Cooper, A. I. Guest control of structure in porous organic cages. *Chem. Commun.* **50**, 9465-9468 (2014).

28 Thallapally, P. K., Dalgarno S. J. & Atwood, J. L. Frustrated organic solids display unexpected gas sorption. *J. Am. Chem. Soc.* **128**, 15060–15061 (2006).

29 MacDowell, D. & Nelson, J. Facile synthesis of a new family of cage molecules. *Tet. Lett.* **29**, 385-386 (1988).

30 Drew, M. G. B., McDowell, D. & Nelson, J. A new ditopic polyaza macrobicyclic ligand: X-ray crystallographic structure determination. *Polyhedron* **7**, 2229-2232 (1988).

31 Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225-11236 (1996).

32 Bojdys, M. J. *et al.* Porous organic cage crystals: characterising the porous crystal surface. *Chem. Commun.* **48**, 11948-11950 (2012).

33 Robeson, L. M. Correlation of separation factor versus permeability for polymeric membranes. *J. Membr. Sci.* 165-185 (1991).

34 Atwood, J. L., Barbour, L. J., Jerga, A. & Schottel, B. L. Guest transport in a nonporous organic solid via dynamic van der Waals cooperativity. *Science* **298**, 1000-1002 (2002).

35 Jin, Y. *et al.* Highly CO_2-selective organic molecular cages: what determines the CO_2 selectivity. *J. Am. Chem. Soc.* 133, 6650-6658 (2011).

36 Jin, Y., Voss, B. A., Noble, R. D. & Zhang, W. A shape-persistent organic molecular cage with high selectivity for the adsorption of CO_2 over N_2. *Angew. Chem. Int. Ed.* 49, 6348-6351, (2010).

37 Schneider, M. W., Oppel, I. M. & Mastalerz, M. Exo-functionalized shape-persistent [2+3] cage compounds: influence of molecular rigidity on formation and permanent porosity. *Chem. Eur. J.* **18**, 4156-4160, (2012).

38 Mastalerz, M., Schneider, M. W., Oppel, I. M. & Presly, O. A Salicylbisimine cage compound with high surface area and selective CO_2/CH_4 adsorption. *Angew. Chem. Int. Ed.* 50, 1046-1051 (2011).

39 Liu, M. *et al.* Acid- and base-stable porous organic cages: shape persistence and pH stability via post-synthetic "tying" of a flexible amine cage. *J. Am. Chem. Soc.* **136**, 7583-7586 (2014).

40 Jones, W., Theocharis, C. R., Thomas, J. M. & Desiraju, G. R. Structural mimicry and the photoreactivity of organic solids. *Chem. Commun.* 1443-1444 (1983).

41 Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **3**, 935-949 (2004).

42 Lewis, D. W., Willock, D. J., Catlow, C. R. A., Thomas, J. M. & Hutchings, G. J. De novo design of structure-directing agents for the synthesis of microporous solids. *Nature* **382**, 604-606 (1996).

43 Pophale, R., Daeyaert, F. & Deem, M. W. Computational prediction of chemically synthesizable organic structure directing agents for zeolites. *J. Mat. Chem. A* **1**, 6750-6760 (2013).

44 Price, S. L. Computed crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. *Acc. Chem. Res.* **42**, 117-126, (2008).

45 Nowell, H., Barnett, S. A., Christensen, K. E., Teat, S. J. & Allan, D. R. I19, the small-molecule single-crystal diffraction beamline at Diamond Light Source. *J. Synchr. Rad.* **19**, 435-441 (2012).

46 Sheldrick, G. M. SADABS. *University of Göttingen, Germany* (2008).

47 Sheldrick, G. M. A short history of SHELX. *Acta Cryst. Sec. A* **64**, 112-122 (2008).

48 Thompson, S. P. *et al.* Fast X-ray powder diffraction on I11 at Diamond. *J. Synchr. Rad.* **18**, 637-648 (2011).

49 Parker, J. E., Potter, J., Thompson, S. P., Lennie, A. R. & Tang, C. C. In situ gas supply system on the powder diffraction beamline I11 at Diamond Light Source. *Mater. Sci. Forum* **706 - 709**, 1707-1712 (2012).

50 TOPAS academic version 4.1 *Coelho Software, Brisbane, Australia* (2007).

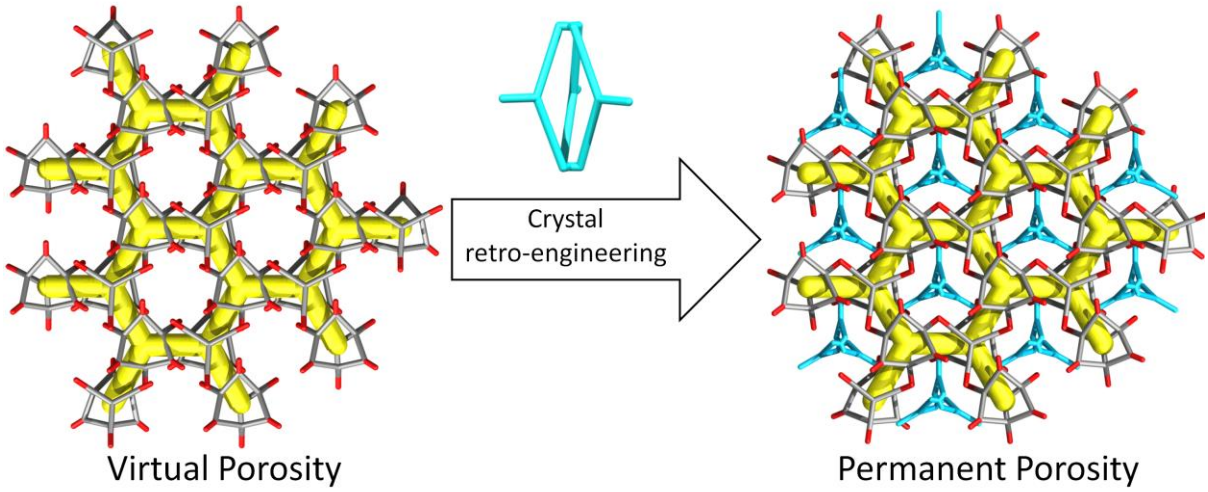


Table of contents graphic: Trapping Virtual Pores by Crystal Retro-Engineering